The Application of Machine Learning in Alzheimer's Disease

Tianhe Yang *

Department of Physics and Astronomy, 525 Northwestern Avenue, Purdue University, 47907-02, West Lafayette

* Corresponding Author Email: yang2184@purdue.edu

Abstract. While mentioning Alzheimer's disease, it's horrible elderly characterized by the loss of important memories and fundamental cognitive disability. Currently, machine learning has been widely applied in the early analysis and recognition of Alzheimer's disease. This paper explores the predictive methods for Alzheimer's disease, including the application of machine learning techniques in this field. The paper compares the traditional MoCa assessment method with the machine learning-based approach, highlighting the latter's advantages in diagnostic accuracy and efficiency. It provides a detailed analysis of the main components of the machine learning predictive method, introducing their functions, advantages, and disadvantages in prediction. Subsequently, through multiple research case studies, the effectiveness and advantages of machine learning methods are demonstrated, showing superiority in both accuracy and universality. Finally, the paper emphasizes the exploration of the causation of Alzheimer's disease and proposes future directions for improvement in research, the importance of interdisciplinary collaboration, and prospects for future research.

Keywords: Machine Learning, Alzheimer, Application.

1. Introduction

Alzheimer's disease, also known as senile dementia, is a neurodegenerative disease primarily affecting the elderly. Symptoms of Alzheimer's disease is marked by memory chaos and cognitive disability, leading to a gradual loss of self-care ability, eventual complete dependence on others, and death [1]. Unlike common diseases, the progression of Alzheimer's is slow but continually worsens over time. According to data provided by the Alzheimer's Association, the number of Americans over 65 with Alzheimer's will exceed 6.7 million in 2023 [2]. Additionally, complete medical research on Alzheimer's disease has not yet been accomplished. Predicting Alzheimer's is crucial. Understanding which populations are at higher risk and conducting early screenings and treatments are essential efforts in positively impacting Alzheimer's management. Even though predicting Alzheimer's is exceptionally challenging, treatments in the early stages are often more effective and less damaging than those in later stages. Therefore, making predictions is a crucial task [3].

Currently, machine learning is a common method used in the prediction of Alzheimer's disease. Many research teams have utilized machine learning techniques to achieve highly accurate predictions for Alzheimer's. Kavitha and her team, using various machine learning methods, achieved an 83% prediction accuracy in a study involving 150 patients [4]. In another study in 2022, Vasco and her team achieved a 90.6% accuracy rate using machine learning methods on 570 patients [5].

This paper will introduce the machine learning prediction methods for Alzheimer's disease and provide specific examples to illustrate why machine learning is an effective approach for predicting Alzheimer's. It will also highlight the advantages of machine learning over traditional questionnaire-based prediction methods, such as the MoCa (Montreal Cognitive Assessment) approach.

2. Traditional Method

2.1. Etiology of Alzheimer's Disease

Unfortunately, the causes of Alzheimer's disease are still unclear, however. Alzheimer's disease is currently considered a neurodegenerative disease, with over 70% of its risk factors believed to be

related to genetics [6]. Age is also a significant factor in Alzheimer's disease, with most cases occurring in the elderly. Additionally, psychological and social environments (such as educational level), lifestyle, and other factors are considered major contributors to Alzheimer's.

However, there are likely other factors that play a significant role in the onset of Alzheimer's, but its complexity has made it difficult to identify them clearly, adding considerable challenges to its exploration and research.

2.2. MoCa Approach

Long before the advent of machine learning, attempts were made to predict Alzheimer's disease. The Montreal Cognitive Assessment (MoCA) is an assessment based on various cognitive abilities such as attention, executive function, memory, and language. It can often identify early symptoms of Alzheimer's disease.

However, this predictive method is not entirely reliable. Daniel and his team conducted an assessment using MoCA on over 350 patients in 2015. Despite having a high specificity of 94% (few false positives), the low specificity rate reached 60%. This means that 40% of those not suffering from dementia were mistakenly identified as patients [7]. In summary, this "questionnaire-based" prediction method does not have the high accuracy required to fully assess whether an individual has Alzheimer's disease. Moreover, this questionnaire-based approach is not feasible for every elderly person, making it difficult for large-scale predictions and more suitable for small-scale screening or targeted identification.

3. Machine Learning Approach

In summary, due to the complex causes of Alzheimer's disease, traditional methods, such as the MoCA assessment, cannot provide satisfactory results. Therefore, with the advent of the era of big data, the prediction of Alzheimer's disease by the method of machine learning would overcome the drawbacks of traditional methods.

3.1. OASIS database

The Open Access Series of Imaging Studies (OASIS) is a significant database for Alzheimer's disease prediction. It is a publicly available neuroimaging dataset consisting of brain MRI scan results. These datasets include brain scan images from individuals of different ages and cognitive levels, along with clinical assessment results and other relevant information, such as age, gender, and education level. All of this data is instrumental in researching and predicting Alzheimer's disease, forming a crucial part of the study [8].

Many machines learning-based predictions of Alzheimer's disease draw their data from OASIS, making this database a valuable resource for Alzheimer's research. The methods mentioned later in the text also utilize the OASIS database.

3.2. Machine Learning and Forecasting

Typically, the prediction of Alzheimer's disease involves more than one machine-learning method. Most research approaches include decision trees, random forests, SVM (Support Vector Machine), XGBoost, and voting. This paper will explain using the method employed by Kavitha and her team as an example [3]. Each method has its unique application scenarios, and combining them achieves the best effect in machine learning predictions.

A decision tree is a data model that assists in data classification and regression. It works by mimicking the human decision-making process. Constructing a decision tree is relatively simple and doesn't require extensive data preprocessing. By using the OASIS database with minimal preprocessing to build a decision tree, doctors can distinguish Alzheimer's patients and conduct basic risk assessments. However, a single decision tree can lead to overfitting.

The emergence of random forests effectively resolves the overfitting problem that may occur with decision trees. A random forest includes more than one decision tree and is a type of ensemble learning method, reducing overfitting and increasing prediction accuracy.

The application of XGBoost can also reduce the errors of single decision tree predictions, resulting in higher accuracy. Moreover, XGBoost can classify the symptoms of Alzheimer's disease, aiding doctors in personalized treatment planning.

The use of Support Vector Machine (SVM) is another significant branch in machine learning (ML) for Alzheimer's disease prediction. This method, ensuring different types of data are effectively segmented, helps doctors understand the stage of Alzheimer's disease a patient is in, greatly reducing the workload in designing medical plans.

Voting combines aspects of the aforementioned machine learning methods to make judgments with smaller errors, avoiding potential mistakes that can occur when the human brain processes complex data.

Kavitha's team combined these machine learning methods and achieved an 83% prediction accuracy in their study of 150 patients, demonstrating the significant medical relevance of ML predictions [3]. In another prediction study, Diogo and his team [4] also used machine learning methods and make predictions of Alzheimer's disease with 570 patients, achieving similarly satisfactory results. Using the OASIS database for predictions led to a 90.6% accurate diagnosis rate, showing tremendous clinical potential.

4. Conclusion

This paper discusses the causes of Alzheimer's disease and compares the traditional MoCa method with machine learning approaches. After analyzing how both methods operate and their actual success rates in prediction, it concludes that machine learning is a more efficient and accurate method for predicting Alzheimer's disease. This study encouragingly observes significant progress in modern technology across various fields, particularly in understanding Alzheimer's disease. Current technologies are gradually improving the predictive ability for this disease, which is crucial for doctors to identify potential patient groups early and intervene timely.

However, the pathogenic mechanism of Alzheimer's disease remains not fully revealed. Although research indicates that familial genetics and psychological states may be related to the disease's progression, the specific roles of these factors and their interactions with other potential causes are unclear. This uncertainty presents significant challenges in predicting and preventing the disease. Additionally, the accuracy of disease prediction using machine learning is limited by the available data and algorithm constraints. Particularly, models based on decision tree algorithms may be affected by factors not fully understood, leading to potential collinearity problems in the models and affecting prediction accuracy.

In summary, Alzheimer's disease is a complex multidisciplinary problem that requires integrating knowledge and techniques from anatomy, psychology, neuroscience, and data science to understand the disease mechanism more comprehensively and improve strategies for prediction, prevention, and treatment. Facing this global challenge, interdisciplinary collaboration and innovative research are crucial to providing more effective help and support for patients.

References

[1] American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV-TR 4th Edition Text Revision. Washington, DC: American Psychiatric Association. 2000. ISBN 0 - 89042 - 025 - 4.

[2] Alzheimer's Association. (n.d.). Alzheimer's disease facts and figures.2023. Retrieved from https://www.alz.org/alzheimers-dementia/facts-figures.

[3] Rasmussen, J., & Langerman, H. < p> Alzheimer’s Disease – Why We Need Early Diagnosis</p> In Degenerative Neurological and Neuromuscular Disease: 2019, 9, 123 – 130.

[4] Kavitha, C., Mani, V., Srividhya, S. R., Khalaf, O. I., & Tavera Romero, C. A. Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models. In Frontiers in Public Health, 2022, 10.

[5] Diogo, V.S., Ferreira, H.A., Prata, D. et al. Early diagnosis of Alzheimer's disease using machine learning: a multi-diagnostic, generalizable approach. Alz Res Therapy 2022,14, 107.

[6] Ballard, C., Gauthier, S., Corbett, A., Brayne, C., Aarsland, D., & Jones, E. Alzheimer's disease. The Lancet, 2011, 377 (9770), 1019 – 1031.

[7] Davis, D. H., Creavin, S. T., Yip, J. L., Noel-Storr, A. H., Brayne, C., & Cullum, S. Montreal Cognitive Assessment for the detection of dementia. In Cochrane Database of Systematic Reviews, 2021, 2021 (7).

[8] Open Access Series of Imaging Studies (OASIS): Cross-Sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. Marcus, DS, Wang, TH, Parker, J, Csernansky, JG, Morris, JC, Buckner, RL. Journal of Cognitive Neuroscience, 19, 1498 - 1507.